



ITS
Institut
Teknologi
Sepuluh Nopember

TUGAS AKHIR - KS141501

ANALISIS FAKTOR- FAKTOR YANG MEMPENGARUHI KESEMBUHAN DAN KEMATIAN PADA PASIEN STROKE DENGAN MENGUNAKAN ASSOCIATION RULE MINING

ERINA SISKA DEWI 5210 100 151

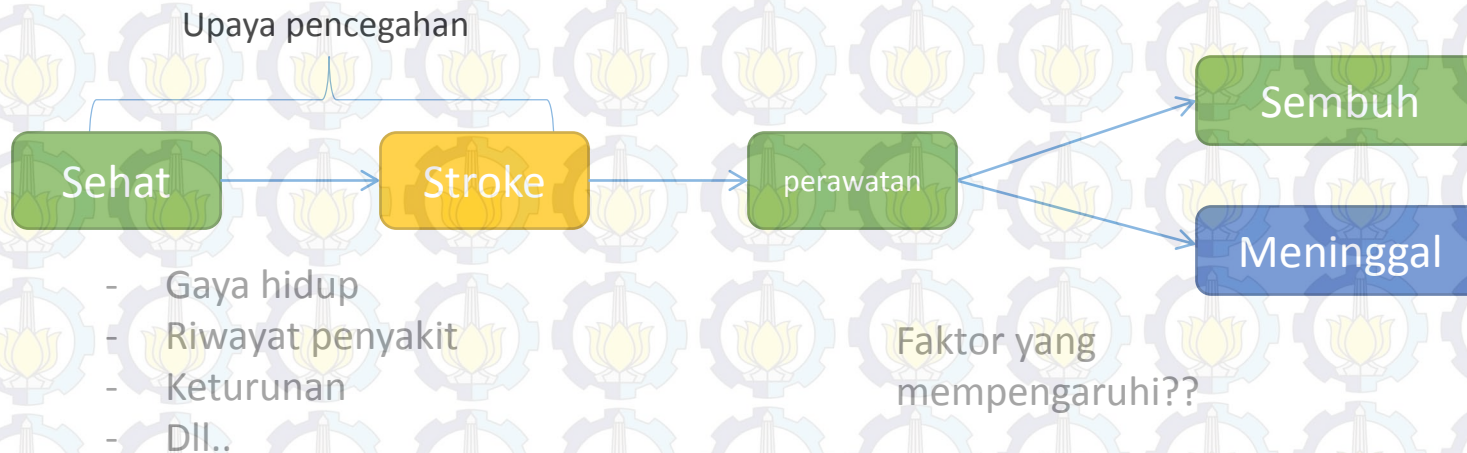
Dosen Pembimbing
Wiwik Anggraeni, S.Si, M.Kom.

Renny Pradina Kusumawardani, S.T., M.T.

Latar Belakang

- 8,3 per 1000 penduduk (2007) menjadi 12,1 per 1000 penduduk (2013) (Badan Penelitian dan Pengembangan Kesehatan, 2013)

Penyakit stroke menjadi penyebab kematian terbanyak ketiga



Association Rule Mining - Apriori



Perumusan Masalah

faktor – faktor apa saja yang mempengaruhi kesembuhan dan kematian pada pasien stroke???

Batasan Masalah

1. Tugas akhir ini menggunakan studi kasus pasien stroke yang dirawat di rumah sakit, data terdiri dari 19.435 pasien yang berasal dari berbagai rumah sakit yang berasal dari 43 negara (data mengenai detail rumah sakit yang terlibat dirahasiakan).
2. Data diambil dari database *International Stroke Trial*. Data pasien yang akan digunakan merupakan data pasien stroke mulai dari tahun 1991 sampai dengan tahun 1996 (5 tahun).
3. Data yang digunakan berisi data pasien yang dipantau masa perawatannya selama 14 hari dan mendapatkan pengawasan selama 6 bulan setelah perawatan di rumah sakit.



Tujuan Tugas Akhir

Mengetahui faktor – faktor yang mempengaruhi kesembuhan dan kematian pasien stroke sehingga dapat digunakan untuk membantu penanganan pasien stroke agar lebih baik, dalam arti mengurangi peluang kematian dan meningkatkan peluang kesembuhan dengan mempertimbangkan faktor – faktor yang berpengaruh terhadap keduanya



Manfaat Tugas Akhir

1. Hasil dari tugas akhir ini diharapkan dapat membantu bidang kedokteran dalam proses penanganan pasien stroke yang tepat.
2. Sebagai referensi untuk mengetahui bagaimana cara menerapkan *association rule mining* dan algoritma apriori.
3. Sebagai masyarakat dapat mengetahui informasi mengenai faktor – faktor yang berpengaruh pada kesembuhan dan kematian pasien stroke, sehingga dapat melakukan tindakan preventif dan tepat sesuai dengan kapasitas yang dimiliki.

Penyakit Stroke

Pembuluh darah pecah/tersumbat

Supply oksigen terganggu

Stroke

Stroke ringan

Stroke iskemik

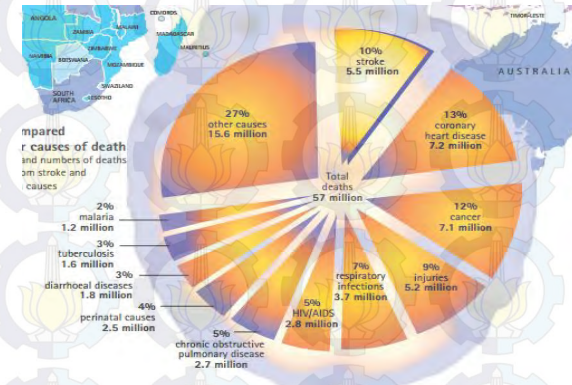
Stroke hemoragik

Trombotik

Embolik

Intraserebral

Subarachnoid





Penyakit Stroke

Penyebab :

Rambe (2006) membagi faktor tersebut menjadi dua yaitu faktor genetik dan gaya hidup.

- Faktor genetik (tidak bisa dimodifikasi) : Usia, ras, jenis kelamin, riwayat penyakit stroke dalam keluarga serta sejarah penyakit stroke yang pernah diderita.
- Gaya hidup : Riwayat penyakit darah tinggi, diabetes mellitus, merokok, hyperlipidemia serta konsumsi makanan dan minuman (bahan pengawet, alkohol, sintetis, dan lain – lain)

Association Rule Mining

Association rule mining merupakan salah satu dari prosedur data mining yang banyak digunakan. Berfungsi untuk mencari pola yang sering muncul diantara banyak transaksi atau untuk menemukan *rule* atau aturan antara suatu kombinasi item

Rule berbentuk $X \rightarrow Y$, dimana X dan Y dapat berisi lebih dari satu item. X dan Y merupakan *disjoint itemset* ($X \cap Y = \emptyset$). Rule tersebut menggambarkan bahwa Y akan terjadi ketika X terjadi

Association rule ini akan menghasilkan *rules* yang menentukan seberapa besar hubungan atau pengaruh antara X dan Y.

Rule yang terbentuk harus memenuhi batas ambang minimum support dan confidence

Association Rule Mining

Support adalah presentase dari transaksi yang mengandung itemset (X dan Y).

$$\text{Support}(X \rightarrow Y) = \frac{\text{Jumlah transaksi yang mengandung } X \text{ dan } Y}{\text{Jumlah transaksi keseluruhan}}$$

Confidence merupakan rasio antara jumlah transaksi yang berisi X dan Y dan jumlah transaksi yang berisi X.

$$\text{Confidence}(X \rightarrow Y) = \frac{\text{Jumlah transaksi yang mengandung } X \text{ dan } Y}{\text{Jumlah transaksi yang mengandung } X}$$

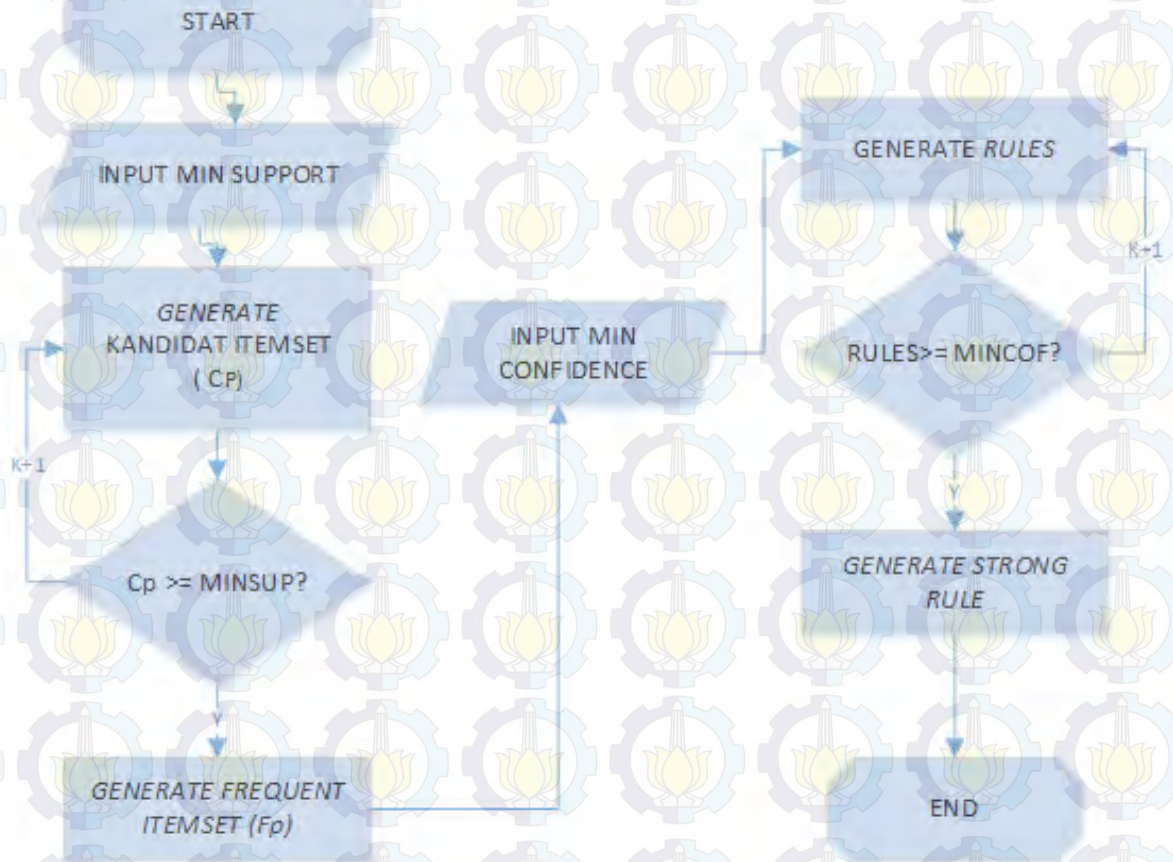
Lift, mengukur tingkat kepentingan dari suatu *rule*.

$$\text{lift}(X \rightarrow Y) = \text{lift}(Y \rightarrow X) = \text{conf}(X \rightarrow Y) / \text{supp}(Y) = \text{conf}(Y \rightarrow X) / \text{supp}(X)$$

Coverage

Mengukur seberapa besar cakupan implementasi *rule* pada keseluruhan data. Dapat dihitung dari nilai support lhs (*left hand side*) / *antecedent*.

Association Rule Mining



Algoritma Apriori

Frequent itemset generation dengan menggunakan algoritma apriori.

Frequent itemset adalah itemset yang memiliki support yang lebih besar atau sama dengan batas ambang minsup (minimum support). (Tan, Steinbach, & Kumar, 2005)

Algoritmanya adalah sebagai berikut :

1. misal $k = 1$
2. bentuk frequent item sets yang terdiri dari k -item
3. ulangi hingga tidak ada lagi frequent item sets yang baru
4. bentuk kandidat item sets dengan panjang $k+1$ dari frequent item sets dengan panjang k
5. buang kandidat item sets yang berisi subset dengan panjang k yang tidak frequent
6. hitung support dari setiap kandidat dengan scanning basisdata
7. eliminasi kandidat yang infrequent

Koefisien Kontingensi

Uji koefisien kontingensi berfungsi untuk mengujur keeratan hubungan antara dua atribut. Uji ini diperlukan dalam tahap *data preprocessing*. Dengan melakukan uji koefisien kontingensi diharapkan semua atribut yang diuji merupakan atribut yang relevan.

Pemangkasan Rule (*pruning*)

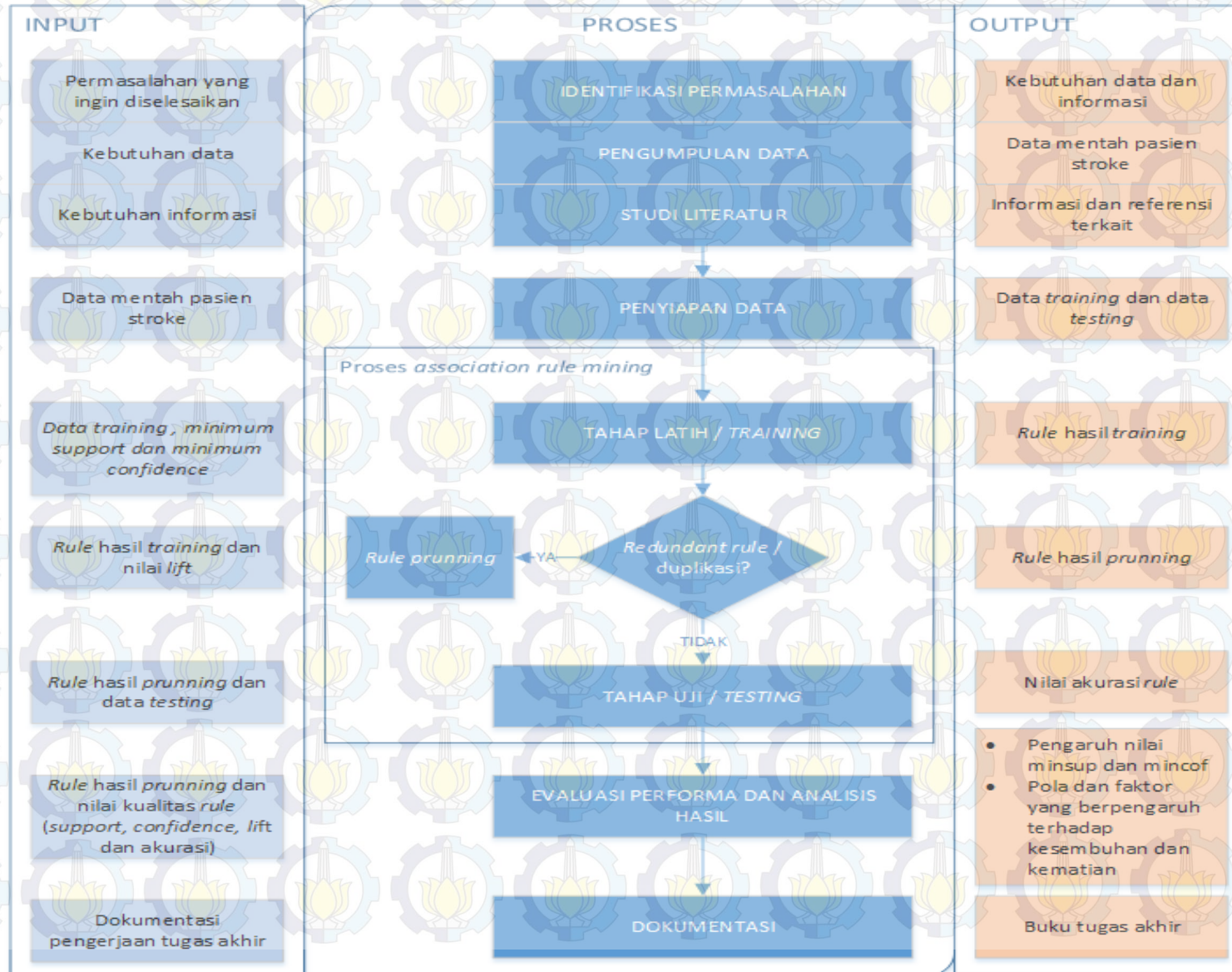
Proses pemangkasan *rule* akan dilakukan setelah proses *generate rule* selesai, bertujuan agar tidak ada rule yang redundan ,

Contoh data keselamatan pasien kapal tenggelam :

	lhs	rhs	support	confidence	lift
1	{Class=2nd, Age=Child}	=> {Survived=Yes}	0.010904134	1.0000000	3.095640
2	{Class=2nd, Sex=Female, Age=Child}	=> {Survived=Yes}	0.005906406	1.0000000	3.095640

Redundan

METODOLOGI





PENGUMPULAN DATA

International Stroke Trial Database

- funded by the UK Medical Research Council
- UK Stroke Association
- European Union BIOMED-1 program
- Bayer USA
- Bayer UK
- National Heart Foundation (Australia)
- Nova Scotia Heart and Stroke Foundation grant(Canada)
- IGA Ministry of Health
- McMaster INCLIN
- Julius Brendel Trust and the Lottery Grants Board
- Norwegian Council on Cardiovascular Disease and Nycomed (for insurance).



Profil data

- 112 Atribut dan 19435 rekam data pasien
- [List atribut](#)

	Number
Ischaemic stroke	17398
Haemorrhagic stroke	599
Indeterminate	992
Not a stroke	420
Uncertain diagnosis	26
Total	19435

Australia Belgium Bulgaria Canada Chile
Czech Republic Denmark Ireland Finland
France Germany Israel Italy Netherlands New Zealand
Norway Poland Portugal Slovenia South Africa
Spain Sri Lanka Sweden
Switzerland Thailand UK
USA Argentina Hong Kong Greece
Georgia Romania Singapore
Turkey Hungary India Japan Latvia
Malaysia Indonesia Brazil Albania Slovak Republic

DATA PREPROCESSING

Penambahan atau pengurangan atribut

Data Diagnosis

Atribut	Koefisien kontingensi	P-Value
RXASP	0.011	0.633
HOURLOCAL	0.093	0.454
MINLOCAL	0.077	0.454
RHEP24	0.02	0.122
RSBP	0.185	0.092
RDATE	0.178	0.092
RASP3	0.081	0.082

Penggabungan Atribut	
DDIAGISC	DIAGNOSIS
DDIAGHA	
DDIAGUN	
DNOSTRK	
DNOSTRKX	



DDIAGISC	DDIAGHA	DDIAGUN	DNOSTRK	DIAGNOSIS
Y	N	N	N	Ischaemic
Y	N	N	N	Ischaemic
Y	N	N	N	Ischaemic
Y	N	N	N	Ischaemic
Y	N	N	N	Ischaemic
Y	N	N	N	Ischaemic
N	N	N	Y	Not s stroke
Y	N	N	N	Ischaemic
Y	N	N	N	Ischaemic
Y	N	N	N	Ischaemic
Y	N	N	N	Ischaemic
Y	N	N	N	Ischaemic
N	Y	N	N	Haemorrhagic
N	N	N	Y	Not a stroke
Y	N	N	N	Ischaemic
N	N	Y	N	Indeterminate

Data Survival

Atribut	Koefisien Kontingensi	P-Value
DDEADD	1.2	1
DRSUNKD	0.67	0.963
DMAJNCHD	0.695	0.947
DRSISCD	0.541	0.938
DPED	0.701	0.93
FU1_COMP	0.763	0.807
FLASTD	0.837	0.685
DSIDED	0.494	0.557
DALIVED	0.271	0.491
DRSHD	0.77	0.461
H14	0.021	0.266
NCB14	0.021	0.201
FU1_RECD	0.04	0.09
NK14	0.26	0.073
DVT14	0.26	0.073
expd6	0.864	0.071
TRAN14	0.33	0.068
FU2_DONE	0.85	0.067
expdd	0.699	0.06
expd14	0.538	0.058

Data Survival

Atribut Dihilangkan	Keterangan
DMAJNCHX	
DSIDEX	Berisi Keterangan atau komentar
DDEADX	
FDEADX	
DIED	Isi sudah tercakup dalam atribut OCCODE yang berisi hasil perawatan pasien
DALIVE	
FRECOVER	
FDENNIS	
DDEAD	
FDEAD	
DEAD1	
DEAD2	
DEAD3	
DEAD4	
DEAD5	Isi sudah terkandung dan sama dengan atribut FDEADC & DDEADC yang berisi penyebab kematian pasien
DEAD6	
DEAD7	
DEAD8	

Selected attribute		
Name: OCCODE		Type: Nominal
Missing: 0 (0%)		Unique: 0 (0%)
Distinct: 5		
No.	Label	Count
1	DEAD	2968
2	DEPENDENT	5535
3	missing status	107
4	NRECOVER	2684
5	RECOVER	2310

Selected attribute		
Name: DCAUSE		Type: Nominal
Missing: 0 (0%)		Unique: 0 (0%)
Distinct: 9		
No.	Label	Count
1	Coronary	277
2	Initial Stroke	1029
3	non vascular	179
4	Other vascular/unknown	370
5	Pneumonia	603
6	Pulmonary embolism	118
7	Recurrent(haemorrhagic/u...	60
8	Recurrent(haemorrhagic/u...	280

Konversi atribut bernilai kontinyu menjadi atribut diskrit

Selected attribute

Name: AGE

Missing: 0 (0%)

Distinct: 6

Type: Nominal

Unique: 0 (0%)

No.	Label	Count
1	'(-inf-52.5]'	1368
2	'(52.5-64.5]'	3270
3	'(64.5-74.5]'	6005
4	'(74.5-82.5]'	5362
5	'(82.5-89.5]'	2940
6	'(89.5-inf)'	490

RCONSC

SEX

AGE

D

M

69

F

M

76

F

F

71

F

M

81

F

M

78

F

M

54

F

F

77

F

M

23

F

M

47

F

M

81

D

M

48

F

F

45

D

F

83

F

F

86

F

M

50

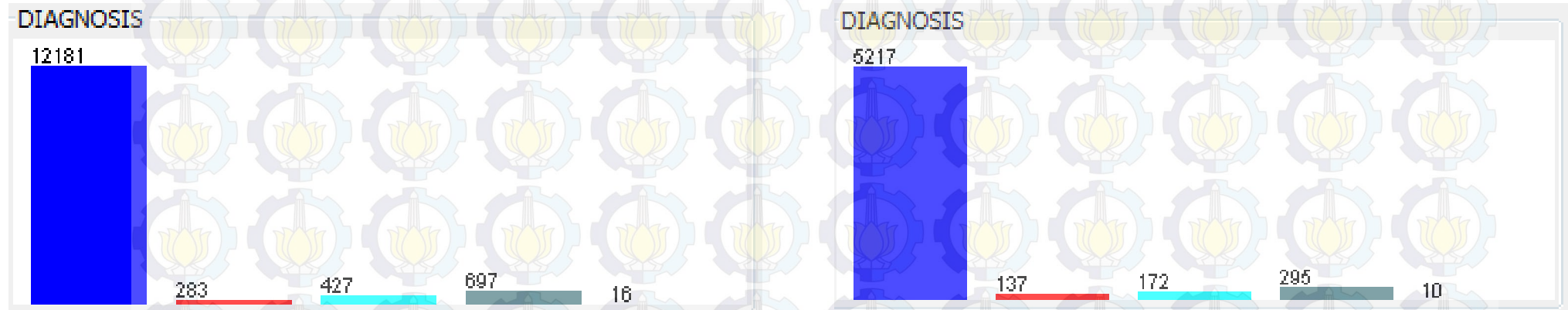


Memilih atribut yang relevan

1. [Data Diagnosis](#)
2. [Data Survival](#)

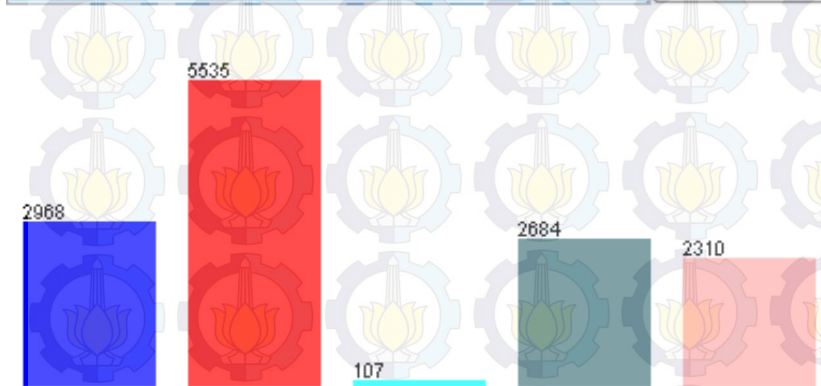
Pembagian data *training* dan data *testing*

Untuk pembagian data *training* dan *testing* akan digunakan filter resample pada WEKA. Data *training* akan dialokasikan sebesar 70% dan 30% untuk data *testing*.



Selected attribute		
Name: OCCCODE		
Missing: 0 (0%)		
Distinct: 5		
Type: Nominal		
Unique: 0 (0%)		
No.	Label	Count
1	DEAD	2968
2	DEPENDENT	5535
3	missing status	107
4	NRECOVER	2684
5	RECOVER	2310

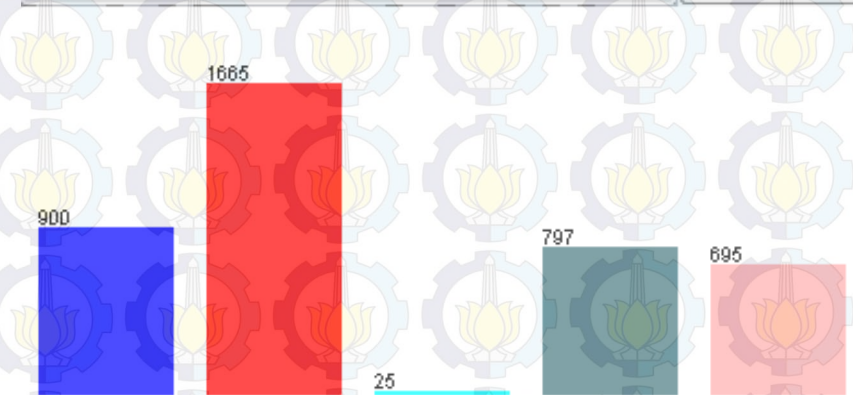
Class: OCCCODE (Nom) Visualize All



OCCCODE data training

Selected attribute		
Name: OCCOD		
Missing: 0 (0%)		
Distinct: 5		
Type: Nominal		
Unique: 0 (0%)		
No.	Label	Count
1	DEAD	900
2	DEPENDENT	1665
3	missing status	25
4	NRECOVER	797
5	RECOVER	695

Class: OCCCODE (Nom) Visualize All



OCCCODE data testing



Penyiapan file sesuai dengan format

Setelah data dipecah menjadi data training dan testing, akan disiapkan data sesuai dengan format yang dapat dibaca oleh aplikasi RStudio. Aplikasi ini memiliki banyak package yang dapat diinstall sendiri oleh user. Salah satu package yang dapat dipakai untuk membaca file adalah package foreign. Berfungsi untuk membaca file dari aplikasi lain seperti excel, minitab, spss dan WEKA. Karena sebelumnya data di praproses menggunakan WEKA, akan dipakai file dengan format .arff untuk diproses didalam RStudio.



PROSES ASSOCIATION RULE MINING

Mencari *Frequent Itemset* dengan menggunakan algoritma apriori

Menghasilkan *frequent itemset* yang memenuhi batas ambang minsup.

Rule Generation

Menghasilkan *rule* yang memenuhi batas ambang *minimum confidence*.



Proses association rule mining

Tahap latih

```
#arules dan arulesviz untuk menjalankan association rule mining
#foreign untuk mengakses file berasal dari luar format rstudio (.Arff)
#load colorspace library
library(arules)
library(arulesviz)
library(foreign)
library(colorspace)

#load data (pilih data latih / training)
survival<-read.arff(file="survivaltraining.arff")

#menjalankan algoritma apriori
#parameter minsup dan minconf dapat disesuaikan
#rhs dapat disesuaikan
rulesurvival<-apriori(survival, parameter= list(support=0.109,confidence=0.130),
  appearance=list(rhs= c("OCCODE=DEAD"),default="lhs"))
```


PROPORSI DATA			Minsup	Mincof
TRAINING	Jumlah	%		
Ischaemic	12181	89.540	0.448	0.537
Haemorrhagic	283	2.080	0.010	0.012
Indeterminate	427	3.139	0.016	0.019
Not a stroke	697	5.123	0.026	0.031
unknown	16	0.118	0.001	0.001
Total	13604	100.000	0.500	0.600
Dead	2068	21.718	0.109	0.130
Recover	1615	16.961	0.085	0.102
Not recover	1887	19.817	0.099	0.119
Dependent	3870	40.643	0.203	0.244
missing status	82	0.861	0.004	0.005
Total	9522	100.000	0.500	0.600

```
#sort rule by lift
rulessurvival.sorted<-sort(rulessurvival, by="lift")
#prune redundant rules
subset.matrix<-is.subset(rulessurvival.sorted, rulessurvival.sorted)
subset.matrix[lower.tri(subset.matrix,diag=T)]<-NA
redundant<-colsums(subset.matrix,na.rm=T)
#menghapus redundant rules
rulessurvival.pruned<-rulessurvival.sorted[!redundant]
#menampilkan
inspect(rulessurvival.pruned)
```

Jika super rule tidak memiliki nilai lift yang lebih besar maka super rule tersebut dianggap rule yang redundan karena tidak menambahkan pengetahuan baru.

Rule redundan dihapus.

```
#menghitung coverage rule
quality(rulessurvival) <- cbind(quality(rulessurvival),
                                coverage = coverage(rulessurvival))
```

OUTPUT

Proses association rule mining

Tahap uji

```
public String[] statKlas(Double DSCH,
    Double DIVH, Double DAP, Double DOAC, Double DGORM,
    Double DCAA, Double DHAEMD, Double DCAREND, Double DTHROMB,
    Double DMAJNCH, Double DCAUSE, Double CMPLASP, Double CMLPHEP,
    Double OCCODE
)
{
    String[] has = new String[16];
    for(int a=0; a < 16; a++){
        has[a] = "False";
    }
    if (DMAJNCH==1&&DCAUSE==9&&CMPLASP==2){
        has[0] = "True";
    }
    if (DCAUSE==9&&CMPLASP==2){
        has[1] = "True";
    }
    if (DSCH==1&&DMAJNCH==1&&DCAUSE==9){
        has[2] = "True";
    }
    if (DIVH==1&&DMAJNCH==1&&DCAUSE==9){
        has[3] = "True";
    }
    if (DSCH==1&&DCAUSE==9){
        has[4] = "True";
    }
    if (DMAJNCH==1&&DCAUSE==9){
        has[5] = "True";
    }
    if (DIVH==1&&DCAUSE==9){
        has[6] = "True";
    }
}
```

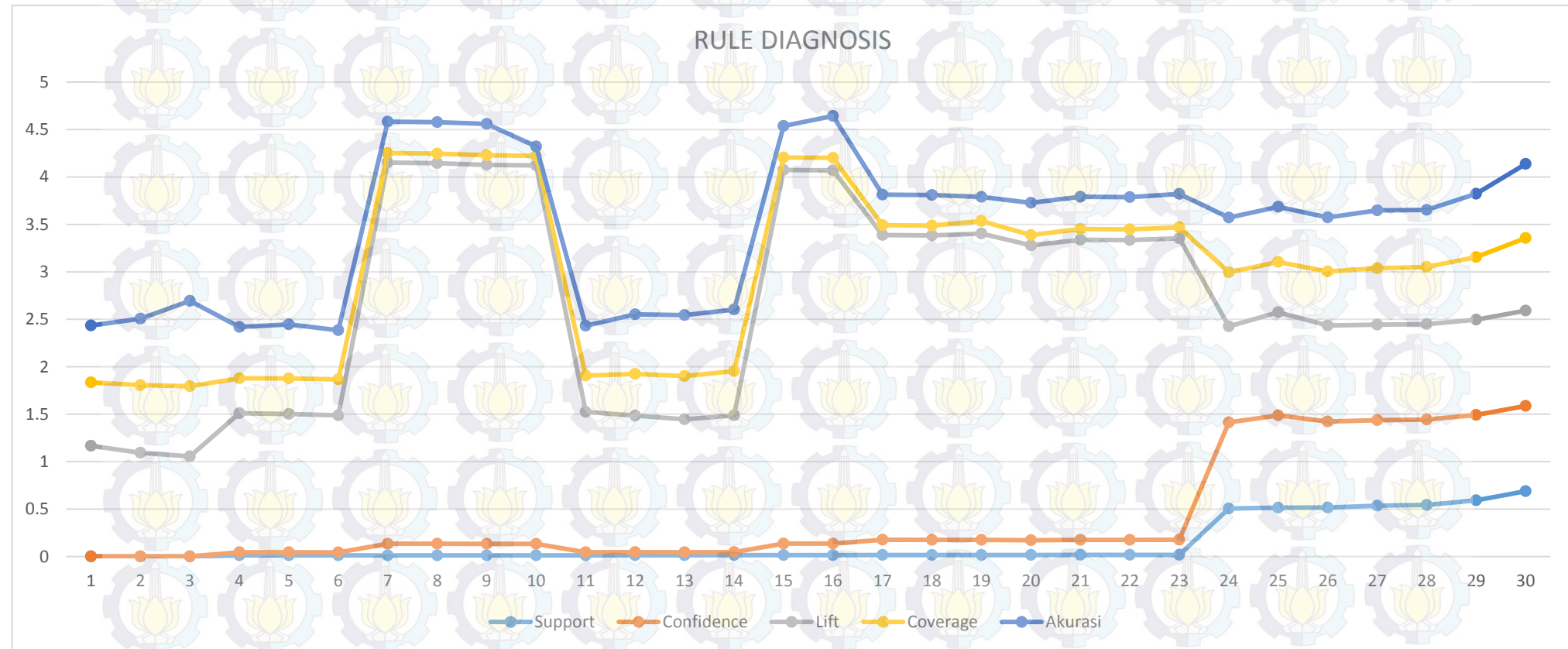
```
public static void main(String[] args) throws IOException {
    MainRun test = new MainRun();
    test.setInputFile("D:/BISMILLAH/Stroke/DATA TESTING/nrecover.xls");
    File output = new File("D:/BISMILLAH/Stroke/DATA TESTING/hasilnrecover.xls");
    test.setData(test.baca());
    int bar = test.data[0].length;

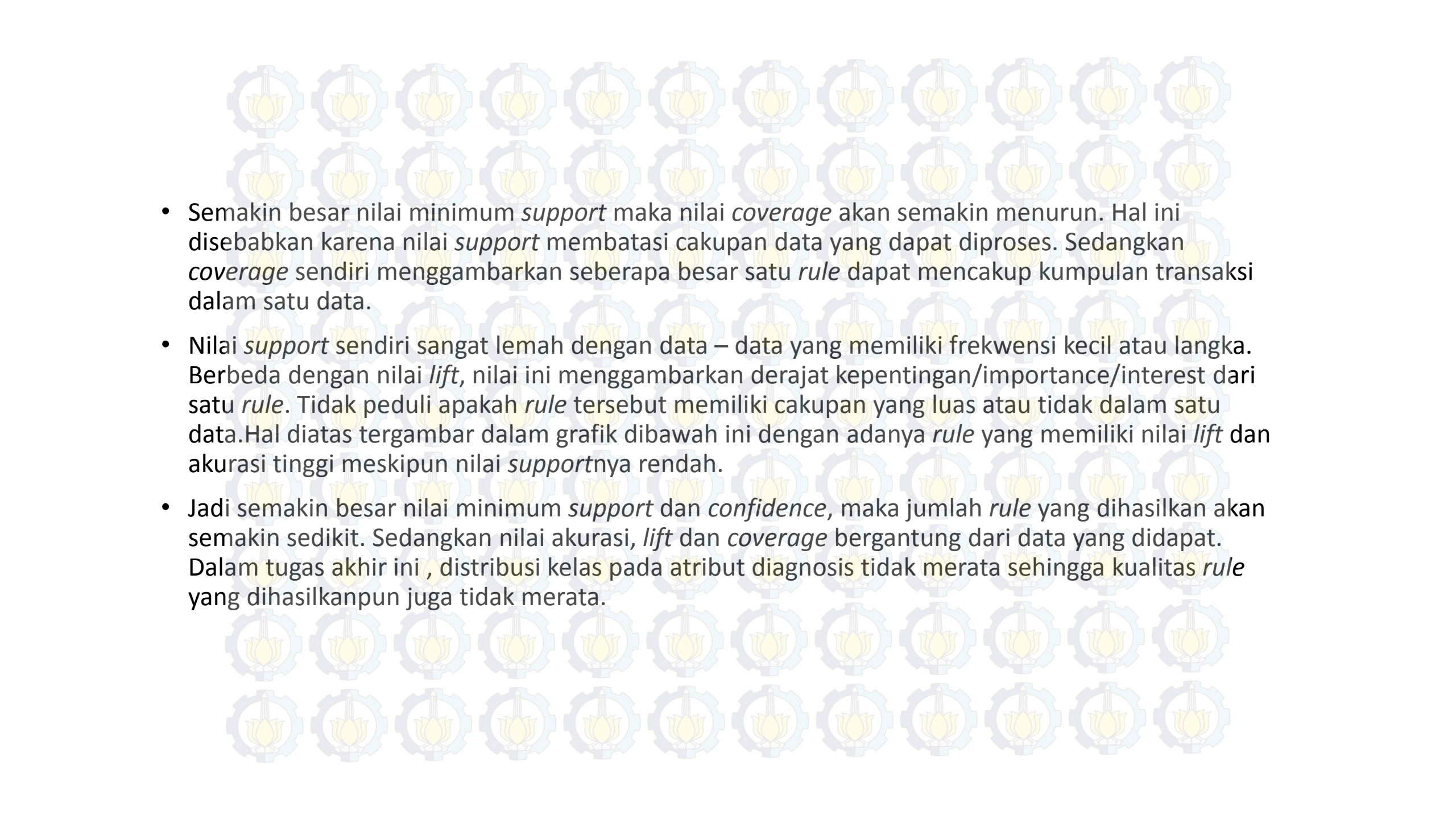
    String[][] has = new String[16][bar];
    String[] hasSem = null;
    recover kp = new recover();
    for(int x=0; x<bar; x++){
        hasSem = kp.statKlas(test.data[0][x], test.data[1][x], test.data[2][x],
            test.data[3][x], test.data[4][x], test.data[5][x], test.data[6][x],
            test.data[7][x], test.data[8][x], test.data[9][x], test.data[10][x],
            test.data[11][x], test.data[12][x], test.data[13][x]);
        for(int y=0; y < 16; y++){
            has[y][x] = hasSem[y]; //memasukkan hasil ke tiap baris
        }
    }

    test.setHasil(has);
    test.tulis(output, test.hasil);
    System.out.println("sukses lihat hasil");
}
```

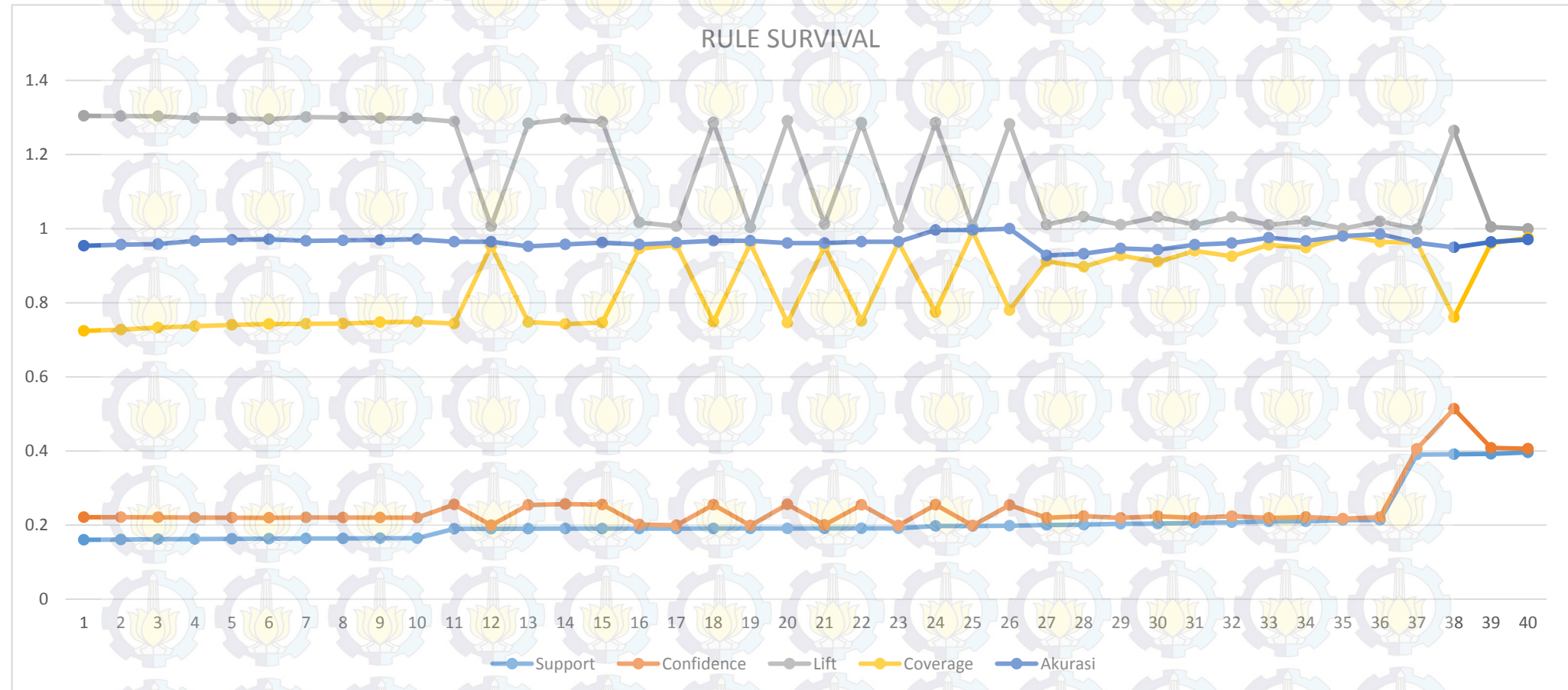
OUTPUT

EVALUASI PERFORMA DAN ANALISIS HASIL



- 
- Semakin besar nilai minimum *support* maka nilai *coverage* akan semakin menurun. Hal ini disebabkan karena nilai *support* membatasi cakupan data yang dapat diproses. Sedangkan *coverage* sendiri menggambarkan seberapa besar satu *rule* dapat mencakup kumpulan transaksi dalam satu data.
 - Nilai *support* sendiri sangat lemah dengan data – data yang memiliki frekwensi kecil atau langka. Berbeda dengan nilai *lift*, nilai ini menggambarkan derajat kepentingan/importance/interest dari satu *rule*. Tidak peduli apakah *rule* tersebut memiliki cakupan yang luas atau tidak dalam satu data. Hal diatas tergambar dalam grafik dibawah ini dengan adanya *rule* yang memiliki nilai *lift* dan akurasi tinggi meskipun nilai *support*nya rendah.
 - Jadi semakin besar nilai minimum *support* dan *confidence*, maka jumlah *rule* yang dihasilkan akan semakin sedikit. Sedangkan nilai akurasi, *lift* dan *coverage* bergantung dari data yang didapat. Dalam tugas akhir ini , distribusi kelas pada atribut diagnosis tidak merata sehingga kualitas *rule* yang dihasilkanpun juga tidak merata.

EVALUASI PERFORMA DAN ANALISIS HASIL



Analisis Rule

No	RULE	Akurasi
1	{RATRIAL=N,RCT=Y} => {DIAGNOSIS=Ischaemic}	58.175%
2	{RDEF1=Y,RDEF2=Y,RDEF7=N} => {DIAGNOSIS=Ischaemic}	56.891%
3	{RCONSC=Y,RDEF2=Y,RDEF7=N}=> {DIAGNOSIS=Ischaemic}	57.658%
4	{RCONSC=F,RDEF6=N} => {DIAGNOSIS=Ischaemic}	60.974%
5	{RDEF1=Y,RDEF3=Y} => {DIAGNOSIS=Ischaemic}	60.073%
6	{RCONSC=F} => {DIAGNOSIS=Ischaemic}	78.072%
7	{RDEF6=N} => {DIAGNOSIS=Ischaemic}	66.763%

- Pasien yang terdiagnosis menderita stroke ischaemic tidak memiliki kelainan detak jantung (RATRIAL) dan sudah melakukan proses CT scan (RCT)
- Pasien terdiagnosis stroke ischaemic dengan gejala adanya deficit/cacat pada wajah(RDEF1), tangan/lengan(RDEF2) dan tidak adanya tanda pendarahan otak (RDEF7)
- Selain itu pasien terdiagnosis stroke ischaemic jika memiliki kesadaran yang kurang/hilang (RCONS), tidak memiliki gejala hilangnya kemampuan visuospasial (RDEF6). Sebagian pasien pada data *testing* juga mengalami deficit pada bagian kaki (RDEF3).

Faktor : **RATRIAL**, RCT, RDEF1,RDEF2, **RDEF7**, RCONS, **RDEF6**, RDEF3

No	RULE	Akurasi
1	{RSLEEP=N,RATRIAL=N,RCT=N,RVISINF=N,RDEF1=Y,RDEF2=Y,RDEF3=Y} => {DIAGNOSIS=Haemorrhagic}	33.140%
2	{RSLEEP=N,RATRIAL=N,RCT=N,RDEF1=Y,RDEF2=Y,RDEF3=Y} => {DIAGNOSIS=Haemorrhagic}	33.140%
3	{RSLEEP=N,RATRIAL=N,RCT=N,RVISINF=N,RDEF1=Y,RDEF3=Y} => {DIAGNOSIS=Haemorrhagic}	33.140%
4	{RSLEEP=N,RATRIAL=N,RCT=N,RDEF1=Y,RDEF3=Y} => {DIAGNOSIS=Haemorrhagic}	9.884%
5	{RSLEEP=N,RATRIAL=N,RCT=N,RVISINF=N,RDEF2=Y,RDEF3=Y} => {DIAGNOSIS=Haemorrhagic}	33.140%
6	{RSLEEP=N,RATRIAL=N,RCT=N,RDEF2=Y,RDEF3=Y} => {DIAGNOSIS=Haemorrhagic}	44.186%

- Penderita penyakit stroke jenis tidak melakukan CT scan sehingga tidak diketahui apakah ada pendarahan otak atau tidak(RVISINF). Gejala yang diderita adalah deficit pada wajah(RDEF1), lengan/tangan (RDEF2) serta bagian kaki (RDEF3).
- Akurasi dari *rule* kurang baik, karena pada saat penggalian *rule*, minimum *support* nilainya sangat kecil karena proporsi data juga sangat kecil.

Faktor : **RSLEEP, RATRIAL,RCT,RVISINF, RDEF1,RDEF2,RDEF3**

No	RULE	Akurasi
1	{RCT=N,RVISINF=N,RDEF1=Y,RDEF2=Y,RDEF4=Y} => {DIAGNOSIS=Indeterminate}	32.203%
2	{RCT=N,RDEF1=Y,RDEF2=Y,RDEF4=Y} => {DIAGNOSIS=Indeterminate}	32.203%
3	{RSLEEP=N,RATRIAL=N,RCT=N,RVISINF=N,RDEF1=Y,RDEF3=Y} => {DIAGNOSIS=Indeterminate}	25.424%
4	{SEX=F,RCT=N,RDEF1=Y} => {DIAGNOSIS=Indeterminate}	35.254%
5	{RCT=N,RVISINF=N,RDEF1=Y,RDEF4=Y} => {DIAGNOSIS=Indeterminate}	33.898%
6	{RCT=N,RDEF1=Y,RDEF4=Y} => {DIAGNOSIS=Indeterminate}	33.898%
7	{RCT=N,RVISINF=N,RDEF2=Y,RDEF3=Y,RDEF4=Y} => {DIAGNOSIS=Indeterminate}	33.898%

Pasien dalam kelompok ini belum bisa ditentukan jenis penyakit strokenya. Tidak melakukan CT scan (RCT) sehingga belum diketahui apakah ada pendarahan otak atau tidak(RVISINF). Namun sudah merasakan gejala – gejala seperti deficit pada bagian tangan dan kaki (RDEF2, RDEF3) serta juga mengalami *dysphasia/language disorder*(RDEF4)

Faktor : **RSLEEP,RCT,RVISINF, RDEF1,RDEF2,RDEF4, SEX**

No	RULE	Akurasi
1	{RATRIAL=N,RVISINF=N,RDEF6=N} => {DIAGNOSIS=Not a stroke}	52.555%
2	{RVISINF=N,RDEF4=N} => {DIAGNOSIS=Not a stroke}	54.015%
3	{RSLEEP=N,RATRIAL=N,RVISINF=N} => {DIAGNOSIS=Not a stroke}	56.934%
4	{RATRIAL=N,RVISINF=N,RDEF5=N} => {DIAGNOSIS=Not a stroke}	51.825%
5	{RVISINF=N,RDEF6=N} => {DIAGNOSIS=Not a stroke}	64.964%
6	{RCONSC=F,RATRIAL=N,RVISINF=N} => {DIAGNOSIS=Not a stroke}	62.774%
7	{RVISINF=N,RDEF5=N} => {DIAGNOSIS=Not a stroke}	64.234%

Pasien digolongkan tidak menderita stroke apabila tidak menderita gejala – gejala stroke kelainan visuospasial, language disorder dan hemianopia/penurunan penglihatan (RDEF6, RDEF4, RDEF5) , tidak ditemukan pendarahan otak (RVISINF) dan tidak ada kelainan detak jantung (RATRIAL).

Faktor : **RATRIAL, RVISINF, RDEF6,RDEF4,RSLEEP,RDEF5,RCONS, SEX**

No	RULE	Akurasi
1	{RVISINF=N} => {DIAGNOSIS=unknown}	60.000%
2	{RSLEEP=N} => {DIAGNOSIS=unknown}	70.000%
3	{RDEF2=Y,RDEF3=Y} => {DIAGNOSIS=unknown}	90.000%

Faktor : **RSLEEP,RCT,RVISINF,**
RDEF1,RDEF2,RDEF4, SEX

Pasien masuk dalam kategori ini jika tidak mengalami pendarahan otak(RVISINF) dan tidak terlihat gejala pada saat pasien tertidur (RSLEEP). Pasien tidak diketahui menderita stroke atau tidak meskipun sudah mengalami gejala deficit lengan dan kaki. Kasus ini dapat disebabkan oleh beberapa pasien yang mengundurkan diri dari perawatan atau pasien menolak tidak bersedia untuk direkam datanya lebih jauh.

No	RULE	Akurasi
1	{DAP=N,DOAC=N,DHAEMD=N} => {OCCODE=DEAD}	93.222%
2	{DAP=N,DOAC=N,DTHROMB=N} => {OCCODE=DEAD}	94.333%
3	{DAP=N,DOAC=N} => {OCCODE=DEAD}	96.111%
4	{DOAC=N,DTHROMB=N} => {OCCODE=DEAD}	96.667%
5	{DOAC=N} => {OCCODE=DEAD}	98.556%
6	{DAP=N,DHAEMD=N,DTHROMB=N} => {OCCODE=DEAD}	92.778%
7	{DAP=N,DHAEMD=N} => {OCCODE=DEAD}	94.667%
8	{DAP=N,DTHROMB=N} => {OCCODE=DEAD}	95.667%
9	{DAP=N} => {OCCODE=DEAD}	97.556%
10	{DTHROMB=N} => {OCCODE=DEAD}	98.000%

DAP : tidak mengkonsumsi obat antiplatelet
DOAC: tidak mengkonsumsi obat antikoagulan lain

DHAEMD: tidak mengalami haemodilution atau penambahan volume plasma darah

DTHROMB : tidak mengalami trombolisis / pemecahan sumbatan darah

Kita dapat menyimpulkan bahwa konsumsi obat antiplatelet dan antikoagulan/pengencer darah sangat mempengaruhi kematian pasien. Dimana obat – obat tersebut berperan penting dalam proses trombolisis.

Antiplatelet : obat penghambat pembentukan penyumbat darah

Antikoagulan: obat pencegah penggumpalan darah

No	Rule	Akurasi
1	{DSCH=N,DMAJNCH=N,DCAUSE=unknown} => {OCCODE=RECOVER}	96.835%
2	{DSCH=N,DTHROMB=N,DCAUSE=unknown} => {OCCODE=RECOVER}	96.978%
3	{DMAJNCH=N,DCAUSE=unknown,CMPLHEP=Y} => {OCCODE=RECOVER}	96.691%
4	{DCAUSE=unknown,CMPLHEP=Y} => {OCCODE=RECOVER}	96.978%
5	{DSCH=N,DCAUSE=unknown} => {OCCODE=RECOVER}	97.122%
6	{DMAJNCH=N,DCAUSE=unknown,CMPLASP=Y} => {OCCODE=RECOVER}	97.122%
7	{DTHROMB=N,DMAJNCH=N,DCAUSE=unknown,CMPLHEP=Y} => {OCCODE=RECOVER}	95.396%
8	{DTHROMB=N,DCAUSE=unknown,CMPLHEP=Y} => {OCCODE=RECOVER}	95.683%
9	{DTHROMB=N,DCAUSE=unknown,CMPLASP=Y} => {OCCODE=RECOVER}	95.827%
10	{DSCH=N,DTHROMB=N,DMAJNCH=N,DCAUSE=unknown} => {OCCODE=RECOVER}	96.691%

DCAUSE : penyebab kematian tidak diketahui/ tidak ada karena pasien tidak meninggal.

DSCH: Pasien tidak mengonsumsi obat heparin melalui suntikan bawah kulit.

DMAJNCH: Pasien tidak mengalami pendarahan otak yang besar/ major.

DTHROMB: Pasien tidak mengalami trombolisis / pemecahan sumbatan pembuluh darah.

CMPLHEP: mematuhi untuk pemakaian / konsumsi heparin

CMPLASP: mematuhi untuk pemakaian / konsumsi aspirin

Disini kita dapat mengetahui pemakaian obat aspirin dan heparin dapat mempengaruhi kesembuhan pasien. Konsumsi obat ini dapat membantu pasien untuk mengurangi resiko pendarahan otak (DMAJNCH).

No	Rule	Akurasi
1	{DIVH=N,DTHROMB=N} => {OCCODE=DEPENDENT}	96.216%
2	{DHAEMD=N,DMAJNCH=N} => {OCCODE=DEPENDENT}	96.396%
3	{DTHROMB=N,DMAJNCH=N} => {OCCODE=DEPENDENT}	97.057%
4	{DTHROMB=N,DMAJNCH=N,DCAUSE=unknown} => {OCCODE=DEPENDENT}	94.955%

Pasien dalam kelompok ini, pada akhir masa pemantauan 6 bulan memiliki kondisi masih bergantung kepada orang lain atau belum mandiri dikarenakan masih ada kecacatan. Mayoritas pasien tidak mendapat suntikan obat heparin didalam pembuluh darah (DIVH) dan tidak mengalami pemecahan sumbatan pembuluh darah (DTHROMB). Selain itu pasien juga tidak mengalami kenaikan volume plasma darah atau *haemodillution* (DHAEMD) dan tidak mengalami pendarahan otak yang *major* (DMAJNCH).

No	Rule	Akurasi
1	{DMAJNCH=N,DCAUSE=unknown,CMPLASP=Y} => {OCCODE=NRECOVER}	95.734%
2	{DCAUSE=unknown,CMPLASP=Y} => {OCCODE=NRECOVER}	96.110%
3	{DSCH=N,DMAJNCH=N,DCAUSE=unknown} => {OCCODE=NRECOVER}	96.487%
4	{DIVH=N,DMAJNCH=N,DCAUSE=unknown} => {OCCODE=NRECOVER}	96.236%
5	{DSCH=N,DCAUSE=unknown} => {OCCODE=NRECOVER}	96.738%
6	{DMAJNCH=N,DCAUSE=unknown} => {OCCODE=NRECOVER}	99.624%
7	{DIVH=N,DCAUSE=unknown} => {OCCODE=NRECOVER}	96.487%
8	{DOAC=N,DCAUSE=unknown} => {OCCODE=NRECOVER}	95.232%
9	{DCAUSE=unknown} => {OCCODE=NRECOVER}	100.000%
10	{DMAJNCH=N,CMPLASP=Y} => {OCCODE=NRECOVER}	95.734%
11	{CMPLASP=Y} => {OCCODE=NRECOVER}	96.110%
12	{DSCH=N,DMAJNCH=N} => {OCCODE=NRECOVER}	96.487%
13	{DIVH=N,DMAJNCH=N} => {OCCODE=NRECOVER}	96.236%
14	{DMAJNCH=N} => {OCCODE=NRECOVER}	99.624%
15	{DSCH=N} => {OCCODE=NRECOVER}	96.738%
16	{DIVH=N} => {OCCODE=NRECOVER}	96.487%

Pasien dalam kelompok ini masih belum sembuh dari penyakit stroke. Faktor – faktor yang mempengaruhi kondisi pasien pada kelompok ini hamper sama dengan pasien pada kategori sebelumnya (*dependent*). Faktor – faktor tersebut antara lain adalah sebagai berikut:

- DMAJNCH : Pasien yang belum sembuh sebagian besar tidak mengalami pendarahan otak yang besar
- DCAUSE : Penyebab kematian tidak ada / tidak diketahui karena pasien belum meninggal
- CMPLASP : Pasien sudah mematuhi konsumsi aspirin.
- DSCH : Pasien tidak menerima suntikan heparin (melalui bawah kulit).
- DIVH : Pasien tidak menerima suntikan heparin yang diberikan langsung ke dalam pembuluh darah.
- DOAC : Pasien tidak mengkonsumsi obat antikoagulan atau penghambat pembekuan darah jenis lain.

Untuk pasien dengan status yang tidak diketahui atau missing status, tidak bisa didapatkan *rule* atau polanya. Hal ini dapat disebabkan oleh data pasien dalam kategori ini yang jumlahnya sangat terbatas. Sehingga dengan minimum *support / confidence* yang kecil sekalipun sistem masih tidak dapat mengeluarkan *rule* atau polanya.

No.	Label	Count
1	DEAD	900
2	DEPENDENT	1665
3	missing status	25
4	NRECOVER	797
5	RECOVER	695

Kesimpulan

- Distribusi kelas dalam atribut dapat mempengaruhi kualitas data. Data yang memiliki distribusi kelas yang merata memiliki kualitas lebih baik. Dibuktikan dengan rata – rata nilai akurasi data survival lebih baik daripada data diagnosis.
- Penentuan minimum *support* dan *confidence* mempengaruhi banyaknya *rule* yang dapat dihasilkan, namun tidak mempengaruhi kualitas data yang lain seperti *lift* dan *coverage*.
- Pada data diagnosis, mayoritas penderita stroke dapat digolongkan dari gejala yang diderita. Namun banyak pasien yang tidak dapat digolongkan karena tidak melakukan proses yang lengkap pada saat diagnosis, contohnya adalah proses CT-Scan.

Kesimpulan

- Akurasi *rule* pada data diagnosis memiliki rata – rata sebesar 49.30% dan data survival sebesar 96.44%.
- Pada data survival, kematian dan kesembuhan pasien sangat dipengaruhi oleh konsumsi obat – obatan yang diberikan pada masa perawatan dan pemantauan, lebih dari 90% pasien pada data *testing* , sembuh dari stroke dengan patuh mengkonsumsi obat (aspirin, heparin, obat antiplatelet / antikoagulan lain. Sehingga tidak timbul pendarahan otak dan tidak pecahnya sumbatan gumpalan darah yang menjadi faktor 90% kematian pasien pada data *testing*.
- Pasien yang belum sembuh dan *dependent* sebagian besar tidak mengalami gejala pendarahan otak, pada data IST didapatkan bahwa pasien kategori ini hanya mengkonsumsi obat aspirin dan tidak mengkonsumsi obat heparin atau obat penghambat pembekuan darah jenis lain.